Navy Personnel Research and Development Center

San Diego, CA 92152-6800 TN 88-53 September 1988

Assessment of Alternative Armed Forces Qualification Test (AFQT) Composite Definitions

NPRDC TN 88-53 September 1988

Assessment of Alternative Armed Forces Qualification Test (AFQT) Composite Definitions

Reynaldo I. Monzon
Paul P. Foley

Reviewed by
Len Kroeker

Released by
John Pass
Director, Personnel Systems Department

Navy Personnel Research and Development Center
San Diego, California 92152

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC TN 88-53	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center 6b. OFFICE SYMBOL (If applicable) 63	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) San Diego, California 92152-6800	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Chief of Naval Operations 8b. OFFICE SYMBOL (If applicable) OP-135	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Washington, DC 20350-2000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 99000N PROJECT NO: O&M,N TASK NO: WORK UNIT ACCESSION NO: WRB-B0112

11. TITLE (Include Security Classification)

Assessment of Alternative Armed Forces Qualification Test (AFQT) Composite Definitions

12. PERSONAL AUTHOR(S)
Monzon, Reynaldo I. and Foley, Paul P.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Note	FROM Oct 85 TO Oct 86	1988 September	27

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Composite comparisons, predictive validity, subgroup score distributions.
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The purpose of this study was to identify the most valid of 17 alternative Armed Forces Qualification Test (AFQT) definitions that would have the least impact upon minorities and women. Concern over the use of speeded tests was the basis for this study. Results indicated that of the 17 alternative AFQT composites, four were found to be promising replacements for the current AFQT composite. After additional analysis, the alternative AFQT composite VE + AR + MK was recommended as the most promising replacement. In comparison to the current AFQT, it had comparable predictive validity and showed a minimum negative impact on females and Blacks in both the AYP and the FY84 Navy Recruit Population (NRP).

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Foley, Paul P.	22b. TELEPHONE (Include Area Code) (619) 553-9663 22c. OFFICE SYMBOL Code 63

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete.

FOREWORD

This research was conducted within MIPR 86-29 Armed Services Vocational Aptitude Battery (ASVAB)--USAF, which was funded by the Office of the Assistant Secretary of Defense (FM&P).

This report is the seventh in a series conducted under this work unit investigating new technology for setting ASVAB standards, developing new test forms, and validating ASVAB against school and job performance. The first (NPRDC Spec. Rep. 83-4) described the development of a deliberate failure key for the Armed Forces Qualification Test (AFQT) portion of the ASVAB, which was designed to identify individuals who, in the event of resumption of the draft, would be motivated to fail the AFQT in order to avoid military service. The second (NPRDC Tech. Rep. 85-19) compared the accuracy of univariate and multivariate correction for range restriction to determine which would be the most accurate to use for ASVAB validation. The third (NPRDC Tech. Rep. 86-17) covered the generalizability of validity data for Navy jobs. The fourth (Journal of Applied Psychology, 1986, 71(4), 641-644) investigated whether the validity of a test is constant throughout the test score range. The fifth (in review) investigated the relationship between mental and educational enlistment standards and job performance of enlisted personnel. The sixth (in review) examined the existence of gender related differences in prediction of school performance on ASVAB and investigated the use of alternate composites to eliminate gender related differences in schools where it was observed. The objectives of the present study were to: (a) assess 17 alternative AFQT composites in relation to the current AFQT, and (b) determine which alternative was the most suitable to replace the current AFQT.

Results are intended for use in policy decisions regarding enlistment standards of Navy military personnel as well as by the research community. This investigation represents a small part of an ongoing effort to assure the optimal use of human resources in the Navy.

JOHN J. PASS
Director, Personnel Systems Department

SUMMARY

Problem

The Armed Services Vocational Aptitude Battery (ASVAB) is a Joint Service Battery used for selection and classification of recruit applicants. When the reference population used for norming the ASVAB was changed from a 1944 metric to a 1980 metric, problems with the speeded tests were detected, which led to questions regarding their use. One speeded test, Numerical Operations (NO), is used in calculating the Armed Forces Qualification Test (AFQT) composite. Because of the recent concerns centered on the ASVAB speeded tests, it was proposed that the NO be removed from the AFQT and that an alternative AFQT composite definition be developed. The service labs produced 17 candidate AFQT definitions, each of which did not contain a speeded test.

Purpose

The purpose of the present study was to identify the most valid of 17 alternative AFQT definitions that would have the least impact upon minorities and women.

Approach

Comparisons of the alternative AFQT composites with the current AFQT were made on the basis of two selected criteria: (1) predictive validity, and (2) subgroup score distributions of minorities and women. Therefore, two methods of analyses were performed. The first method (Method 1), produced validity coefficients between the 17 alternative AFQT definitions (including the current AFQT) and the final school grade (FSG) in 24 Navy class "A" schools. For the second method (Method 2), subgroup score distributions by gender and race across AFQT mental levels were developed for the current AFQT composite and the 17 alternative AFQT definitions. The distributions produced by the alternatives were then compared to the current AFQT distribution in order to assess the impact each alternative would have on gender and race. Two samples were utilized: (1) the American Youth Population (AYP), and (2) the FY84 Navy Recruit Population (NRP).

Results

Method 1. All 17 alternative AFQT composites had validities that were either equal to or greater than the validity for the current AFQT. Thus, it is difficult to determine a viable alternative AFQT composite based solely on the predictive validity criterion.

Method 2. In Sample A, alternative AFQT composites VE+AR+MK, VE+AR, VE+MK, and VE+MK+GS showed relatively little change in score distributions of females and Blacks from the current AFQT. For Sample B, these four composites also showed minimal negative impact for females and Blacks.

Conclusions

Of the 17 alternative AFQT composites, four were found to be promising replacements for the current AFQT composite. In order to determine which of the four would be recommended, additional criteria were considered. These criteria that were employed in a parallel study included: (1) AFQT length, the new AFQT should be long enough to minimize test compromise, and (2) AFQT content; the new AFQT should have a balance of

verbal and math constructs and less weight on the technical construct. Only VE+AR+MK satisfied both of these requirements.

Recommendations

The alternative AFQT composite VE+AR+MK is recommended as the most promising replacement for the current AFQT. Compared to the current AFQT, it had comparable predictive validity and showed a minimum negative impact on females and Blacks in both the AYP and the FY84 NRP. Finally, it satisfied additional criteria and was selected as a promising candidate in two parallel studies.

CONTENTS

LIST OF TABLES

INTRODUCTION

Background

The Armed Services Vocational Aptitude Battery (ASVAB) is a Joint Service Battery used for selection and classification of recruit applicants. The Armed Forces Qualification Test (AFQT) formed of 4 of the 10 subtests; Word Knowledge (WK), Paragraph Comprehension (PC), Arithmetic Reasoning (AR), and one half of Numerical Operations (NO) is used to determine an applicant's eligibility for enlistment. As a result of recent concerns regarding the current definition of the AFQT, the present study examines the use of other combinations of ASVAB subtests as possible alternative AFQT definitions.

Prior to 1980, the AFQT consisted of the WK, AR, and Space Perception (SP) tests. The definition of the AFQT changed with the introduction of ASVAB Forms 8, 9, and 10 in October 1980. The WK test was replaced by Verbal (VE), which is the combination of the WK and PC tests, the SP test was dropped from the AFQT and the ASVAB battery altogether. The NO test was added to the current AFQT to make it a longer and more reliable selection instrument. Because it contains more test items, the current AFQT is believed to discourage coaching of recruit applicants, which would have been easier on the prior and shorter AFQT.

In addition to specifying a minimum score on the AFQT, the Navy uses a table--Success Chances for Recruits Entering the Navy (SCREEN)--to determine eligibility. In the Navy, an applicant's AFQT score is combined with age and level of education to determine the probability of their attriting from boot camp. Only applicants with a high probability of successfully completing the first enlistment are accepted.

Problem

The major problem with the current AFQT composite centers around the speeded test, NO. Initial concerns with the ASVAB speeded tests occurred during the reference scaling of the ASVAB and AFQT to the 1980 American Youth Population (AYP) of 18 to 23 year olds (Maier & Sims, 1982; Sims & Maier, 1983). It was found that the average scores on the NO and the other speeded test, Coding Speed (CS) for the 1980 AYP sample were lower than the scores for recruits and military applicants. Earles, Giuliano, Ree, and Valentine (1983) suggested this discrepancy could have been the result of the differences in the answer sheets used, which in turn affected the time sensitivity of the speeded tests. A study by Wegner and Ree (1985) further verified these results.

Although the scores on the speeded tests were eventually corrected for the 1980 AYP reference population, the issue of sensitivity in the speeded tests remains. For example, in October 1984 during the Initial Operational Testing and Evaluation (IOT&E) of ASVAB forms 11/12/13, it was found that the average scores on the new ASVAB forms were two raw score points lower than the average score on the anchor test (ASVAB form 8A). Further investigation by the Air Force revealed differences in print style and format in the NO subtest between ASVAB forms 11/12/13 and the anchor test, which accounted for this discrepancy. Since NO is used in calculating the AFQT, the accuracy of measured trainability by the AFQT may be affected. Therefore, it was proposed that the NO be removed from the AFQT and that an alternative AFQT composite definition be developed. As a result, the service labs produced 17 candidate AFQT definitions, each of which did not contain a speeded subtest.

Horne (1986), who investigated the 17 proposed definitions, discussed several considerations such as length and composition of the alternatives in selecting a comparable AFQT alternative. In terms of length, the 17 candidate definitions vary in the number of tests and test items each one contains. Horne suggests that the new AFQT definition should include an adequate number of tests and test items for several reasons. First, a longer AFQT reduces the probability of test compromise. Second, variation due to true ability rather than chance becomes greater with longer tests. Finally, raw AFQT scores are converted to percentiles. A longer test would result in fewer percentile score gaps, which means a more accurate transformation of scores.

In terms of what combination of ASVAB subtests should make up the composition of the new AFQT definition, Horne (1986) proposes two standards. First, the definition should have an equal balance of verbal and math constructs because they are both considered important general measures of trainability. Second, the technical construct should have a lesser weight because of its negative impact on females.

Purpose

The purpose of the present study was to: (1) evaluate 17 proposed alternative AFQT definitions with regard to predictive validity and race and gender distributions relative to the current AFQT, and (2) determine which alternative was the most suitable to replace the current AFQT.

APPROACH

Two criteria were used as a basis for comparison between the alternative AFQT definitions and the current AFQT. The first criterion, predictive validity, was chosen because the AFQT is primarily used as a measure of general trainability; thus, any suitable definition should be predictive of training performance. The second criterion consisted of various distributions of minorities and women across AFQT mental levels as defined by the alternative AFQT definitions. This criterion was selected to assess any negative impact that these definitions might have on the recruitment of minorities and women into the Armed Services.

Alternative Definitions

Seventeen alternative AFQT definitions were proposed by the service laboratories as candidate replacements for the current AFQT. Table 1 lists the current AFQT and all 17 alternative AFQT definitions.

Table 1

Current AFQT and Proposed
Alternative AFQT Definitions

Current	= VE+AR+NO/2
ALT-a	= VE+AR+MC
ALT-b	= GS+MK
ALT-c	= VE+AR+MK
ALT-d	= VE+AR+GS
ALT-e	= 2AR+2WK+2PC+MK
ALT-f	= VE+2AR+GS
ALT-g	= VE+MK+MC
ALT-h	= GS+AR+AS
ALT-i	= VE+AR
ALT-j	= VE+AR+GS+AS+MK+MC+EI
ALT-k	= VE+AR+GS+MK
ALT-l	= VE+AR+GS+MK+MC
ALT-m	= VE+AR+GS+MK+MC+EI
ALT-n	= WK+2PC+AR+GS+AS+MK+EI
ALT-o	= VE+AR+GS+AS+MK+EI
ALT-p	= VE+MK
ALT-q	= VE+MK+GS

METHOD 1

Sample

Data were obtained from 7,204 students attending 24 Navy class "A" schools from October 1980 to October 1982. The schools were clustered under one of six Navy selector composite groups. All six selector composite groups and the schools within each cluster are listed in Table 2.

Procedure

Uncorrected validity coefficients were computed for each school by correlating the current AFQT and the alternative AFQT definitions with final school grade (FSG).

Table 2

Schools Grouped by Operational Selector Composite

General Technical (VE+AR)

6001 Quartermaster
6005 Signalman
6125 Mess Management Specialist
6167 Data Processing Technician
6301 Cryptologic Technician (Collection)
6302 Cryptologic Technician (Technical) Non-Morse Basic Prep
6477 Ship's Serviceman

Mechanical (VE+MC+AS)

6025 Gunner's Mate Technician (Nuclear Weapons)
6119 Hull Maintenance Technician (San Francisco)
6120 Hull Maintenance Technician (Philadelphia)
6286 Builder
6292 Equipment Operator

Electronics (AR+MK+EI+GS)

6027 Fire Control Technician (Missile) Phase I
6108 Fire Control Technician (Missile) Phase II
6131 Data System Technician
6146 Strategic Weapons Systems Electronics (Polaris-Poseidon)
6376 Fire Control Technician (Gun) Phase II
6377 Fire Control Technician (Gun) Phase I

Clerical (VE+NO+CS)

6020 Cryptologic Technician (Administrative)

BE/E (AR+2MK+GS)

6036 Torpedoman's Mate (Surface Operator) Basic Class "A" School
6070 Electrician's Mate
6093 Torpedoman's Mate (Technician) Basic Class "A" School
6289 Construction Electrician

MR (AR+MC+AS)

6068 Machinery Repairman

Results and Discussion

Tables A-1 through A-6 in Appendix list the coefficients for each school by class "A" school selector composite clusters. For example, Table A-1 contains validity information for each of the seven schools within the General Technical cluster. The class "A" school course numbers are listed across the top of the page and the number of cases in each are listed beneath. The validities of the General Technical Composite (VE+AR), the current

AFQT percentile score, and the current AFQT raw score are listed next followed by the validities of all 17 alternative AFQT definitions. The final column labeled "AVG" provides the weighted average validities across schools within the cluster. Tables A-2 through A-6 contain similar information for each of the remaining selector composite cluster of schools.

The average weighted validity coefficients across schools for the alternative AFQT definitions range from .33 to .38 in the General Technical cluster. Validity coefficient ranges for the remaining clusters are: .33 to .40, Mechanical; .29 to .40, Electronics; .17 to .28, Clerical; .27 to .36, Basic Electricity and Electronics; and .11 to .31, Machinery Repairman.

Table 3 presents the six selector composite clusters, number of schools within each cluster and the average weighted validities within each cluster for the AFQT percentile, raw AFQT, and 17 AFQT alternatives. The last column contains the average validities across the six selector composite clusters. The average coefficients across clusters for all 17 alternatives are either equal to or higher than the average validity coefficient for the raw AFQT. There are several explanations why the current AFQT would have low predictive validity. First, entrance into the Navy is based on the current AFQT. Therefore, the AFQT would be subject to the greatest amount of range restriction compared to any other composite, thereby reducing its predictive value. Second, the current AFQT is composed of only four ASVAB subtests. The number of subtests used for the alternatives range from three to as many as eight. A composite with eight subtests, as opposed to a composite with only three, would yield higher predictive validity. Finally, the current AFQT has a speeded test (NO) that has been found to be sensitive to various testing conditions. In summary, it is difficult to determine a viable alternative AFQT composite definition based solely on the predictive validity criterion.

Table 3

Average AFQT Validities Across Schools Within Clusters

Cluster N of Schools	GT 7	Mech 5	Elec 6	Cler 1	BE/E 4	MR 1	Avg
AFQT%	.37	.32	.29	.25	.32	.13	.32
RAW AFQT	.36	.32	.29	.25	.32	.14	.31
AFQT Alternatives							
ALT-a	.36	.36	.35	.28	.34	.24	.35
ALT-b	.35	.33	.36	.17	.30	.25	.33
ALT-c	.38	.36	.39	.24	.32	.21	.36
ALT-d	.36	.34	.29	.25	.32	.11	.32
ALT-e	.38	.36	.35	.25	.31	.18	.34
ALT-f	.37	.35	.31	.25	.33	.15	.33
ALT-g	.35	.36	.40	.25	.32	.28	.35
ALT-h	.33	.35	.31	.20	.35	.25	.32
ALT-i	.36	.33	.30	.26	.30	.12	.31
ALT-j	.37	.40	.39	.22	.36	.31	.37
ALT-k	.38	.37	.37	.24	.33	.19	.35
ALT-l	.38	.39	.40	.25	.35	.26	.37
ALT-m	.38	.39	.39	.24	.36	.28	.37
ALT-n	.37	.39	.38	.21	.34	.24	.36
ALT-o	.37	.39	.38	.21	.35	.26	.36
ALT-p	.35	.33	.36	.22	.27	.16	.32
ALT-q	.35	.33	.34	.21	.29	.15	.32

METHOD 2

In Method 2, the impact of adopting a new AFQT definition was evaluated by estimating the number of minorities and women falling into each of the AFQT mental level categories and comparing this with what is provided by the current AFQT. Two samples were utilized: (1) the AYP, and (2) the FY84 Navy Recruit Population (NRP). Table 4 contains the different mental level categories (Maier & Sims, 1986) and the AFQT percentile score boundaries for each.

The AFQT raw scores that define the mental level categories were determined using the 1980 AYP, which was administered ASVAB Form 8A. Table 5 contains the lower raw score boundaries corresponding to each of the six mental level categories of Table 4 based on the distributions of the 17 alternative AFQT definitions in the AYP population. As an example, a raw score of 97 for ALT-a equals an AYP percentile score of 93, which defines the bottom of the AFQT mental level category I. Once raw score mental level boundaries were determined, the percent of males, females, whites, and blacks in each AFQT category was estimated for each of the 17 alternative AFQT definitions.

Table 4

AFQT Mental Level Categories

AFQT Mental Level Category	AFQT Percentile Score
I	93-99
II	65-92
IIIA	50-64
IIIB	31-49
IVA, B, C	10-30
V	1-9

Table 5

Raw Score Equivalent to Mental Level Cut Off Points

Mental Level	I	II	IIIA	IIIB	IV	V
Lower Boundary	93	65	50	31	10	
AFQT	100.0	86.0	78.5	66.0	43.0	
ALT-a	97	81	73	60	39	
ALT-b	46	35	29	24	16	
ALT-c	100	81	72	58	37	
ALT-d	99	83	75	62	40	
ALT-e	176	146	130	106	67	
ALT-f	127	105	93	76	49	
ALT-g	92	76	68	56	37	
ALT-h	72	56	49	40	27	
ALT-i	77	65	59	48	30	
ALT-j	178	145	129	106	71	
ALT-k	122	99	88	72	47	
ALT-l	141	115	102	84	56	
ALT-m	158	129	114	94	62	
ALT-n	170	141	127	105	69	
ALT-o	157	128	115	95	63	
ALT-p	72	60	54	44	28	
ALT-q	94	78	70	59	37	

Sample A--1980 American Youth Population

This sample consists of data collected by the National Opinion Research Center (NORC). It contains 9,173 youths between the ages of 18 and 23 and is statistically weighted to be representative of approximately 21 million American youths in 1980.

Procedure

The percentage distributions of males and females across mental levels for the AFQT and the alternative AFQT definitions were estimated and are listed in Table 6. The last column labeled "Total Diff" is the sum of the differences (absolute difference) between the percent of males and females in each of the six AFQT mental level categories. Utilizing the same procedure, percentage distributions were also computed for Whites and Blacks in the AYP. These distributions are listed in Table 7.

Results and Discussion

As indicated by Table 6, the percent of females is less than the percent of males in categories I and II and greater than males in categories IIIA, IIIB, and IV for the current AFQT. This trend was also found for the 17 proposed alternatives. However, the percent of females versus the percent of males varied in category V.

Inspection of Table 7 shows that the percent of Whites is greater than the percent of Blacks in mental level categories I, II, IIIA, and IIIB and less than Blacks in categories IV and V for the current AFQT. This trend is also consistent with all 17 of the proposed alternatives.

The last column in Tables 6 and 7 labeled "Total Diff" is the sum of the differences (absolute difference) between the percent of males and females (Table 6) and the percent of Whites and Blacks (Table 7) across the mental level categories. As an example, for ALT-a in Table 6, the percent of males in category I is 10.7 percent while females have 2.8 percent indicating a difference of 7.9 percent. The differences in the remaining mental levels for ALT-a are: 7.5 percent, category II; -2.1 percent, category IIIA; -6.4 percent, category IIIB; -6.7 percent, category IV; and -0.2 percent, category V. Thus, the absolute sum of these differences is 30.8 percent. For the present study, the Total Diff score was the measure chosen to assess the impact of the 17 alternative AFQT definitions compared to the current AFQT.

In Table 6, the Total Diff score for the current AFQT is 14.8 percent while the 17 alternative definitions have scores that range from 10.5 to 58.4 percent. Four alternative definitions have Total Diff scores that are comparable to the current AFQT. They are: (1) ALT-p (VE+MK) 10.5 percent; (2) ALT-q (VE+MK+GS) 15.6 percent; (3) ALT-i (VE+AR) 16.7 percent; and (4) ALT-c (VE+AR+MK) 16.8 percent. Adoption of any of these four alternatives would mean relatively little change from the current AFQT.

In terms of race, Table 7 shows a Total Diff score of 95.3 for the current AFQT. The Total Diff scores for the alternative definitions range from 84.9 to 104.8 percent. The following alternatives had either lower Diff scores or, were at the most, two percentage points higher than the Diff score for the AFQT: (1) ALT-b (GS+MK) 84.9 percent; (2) ALT-p (VE+MK) 88.7 percent; (3) ALT-c (VE+AR+MK) 92.2 percent; (4) ALT-q (VE+MK+GS) 93.2 percent; (5) ALT-k (VE+AR+GS+MK) 93.3 percent; (6) ALT-e (2AR+2MK+2PC+MK) 93.4 percent; (7) ALT-g (VE+MK+MC) 94.4 percent; and (8) ALT-i (VE+AR) 95.2 percent.

Table 6

Percentage Distributions by Gender Across Mental Level
for American Youth Population

	Mental Levels												
	I		II		IIIA		IIIB		IV		V		
Definitions	M	F	M	F	M	F	M	F	M	F	M	F	Total Diff
AFQT	7.6	5.7	29.6	26.3	14.1	16.8	17.6	20.4	20.0	21.9	11.1	8.8	14.8
ALT-a	10.7	2.8	32.5	25.0	13.4	15.5	16.1	22.5	17.7	24.4	9.6	9.8	30.8
ALT-b	11.4	3.4	27.5	26.0	17.2	16.8	15.2	18.5	20.0	25.1	8.6	10.2	19.9
ALT-c	8.3	5.0	30.6	26.8	13.4	14.5	17.4	21.8	19.8	22.7	10.5	9.2	16.8
ALT-d	9.6	3.9	30.8	26.4	14.4	16.2	16.2	21.4	18.7	22.5	10.3	9.5	21.7
ALT-e	8.7	5.2	30.1	25.8	14.1	16.0	16.8	21.4	19.6	22.4	10.7	9.2	18.6
ALT-f	10.9	4.2	29.8	24.7	13.7	16.2	16.4	21.3	19.4	23.5	9.7	10.0	23.6
ALT-g	9.7	2.5	31.8	26.3	14.4	15.9	16.5	22.1	16.9	23.4	10.7	9.9	27.1
ALT-h	11.9	0.6	37.9	20.1	14.5	16.1	13.8	23.5	14.4	27.8	7.4	12.0	58.4
ALT-i	8.4	5.0	31.0	27.4	13.4	14.7	17.4	22.1	19.5	21.8	10.3	8.9	16.7
ALT-j	12.5	1.1	33.4	22.0	14.1	16.4	15.4	23.5	16.0	25.7	8.6	11.3	45.6
ALT-k	9.5	3.8	30.5	26.9	13.8	15.9	16.8	20.8	19.2	22.8	10.2	9.8	19.4
ALT-l	11.6	3.0	30.8	24.9	14.0	16.2	15.7	21.2	18.2	23.9	9.7	10.7	28.9
ALT-m	11.8	1.7	31.5	24.2	14.4	16.7	15.3	21.6	17.9	25.2	9.1	10.6	34.8
ALT-n	13.0	1.6	31.3	23.9	13.9	15.9	15.5	23.4	17.0	24.2	9.3	11.0	37.6
ALT-o	12.3	1.2	33.0	24.0	13.8	15.5	14.9	22.8	16.9	25.3	9.2	11.2	40.1
ALT-p	7.6	4.7	28.7	28.4	14.4	15.9	19.4	22.0	18.7	19.8	11.3	9.2	10.5
ALT-q	9.4	3.8	28.8	27.9	14.9	15.9	16.3	20.0	20.2	23.3	10.4	9.1	15.6

Table 7

Percentage Distributions by Race Across Mental Level
for American Youth Population

	Mental Levels												
	I		II		IIIA		IIIB		IV		V		
Definitions	W	B	W	B	W	B	W	B	W	B	W	B	Total Diff
AFQT	7.7	0.4	31.5	5.9	17.0	5.6	19.4	16.0	18.1	38.8	6.4	33.3	95.3
ALT-a	7.8	0.4	32.6	4.7	15.7	6.2	19.9	15.1	18.0	39.9	5.9	33.6	99.2
ALT-b	8.5	1.0	29.9	7.2	18.3	9.0	17.2	14.2	19.9	39.1	6.2	29.4	84.9
ALT-c	7.6	0.4	32.1	7.3	15.3	5.6	20.2	15.8	18.4	38.8	6.4	32.1	92.2
ALT-d	7.8	0.4	32.2	6.2	16.7	6.2	19.4	15.0	17.7	38.9	6.2	33.3	96.6
ALT-e	8.0	0.5	31.4	6.6	16.5	6.1	19.6	15.6	18.2	38.5	6.3	32.7	93.4
ALT-f	8.8	0.6	30.7	5.5	16.4	5.9	19.5	14.4	18.5	40.1	6.1	33.5	98.0
ALT-g	7.1	0.4	32.7	6.3	16.4	6.7	19.9	15.3	17.5	36.5	6.4	34.9	94.9
ALT-h	7.3	0.5	33.1	4.1	16.7	6.2	19.4	13.3	17.8	41.4	5.7	34.5	104.8
ALT-i	7.7	0.6	32.8	6.2	15.4	5.5	20.3	16.3	17.7	39.3	6.1	32.1	95.2
ALT-j	7.9	0.6	31.5	4.5	16.7	6.1	20.3	13.3	17.8	39.5	5.8	36.0	103.8
ALT-k	7.6	0.6	32.2	6.8	16.2	6.1	19.3	15.3	18.3	37.5	6.3	33.9	93.3
ALT-l	8.5	0.6	31.4	5.5	16.4	6.5	19.1	14.2	18.3	38.2	6.3	35.1	97.3
ALT-m	7.8	0.6	31.6	4.9	17.0	6.7	19.1	13.9	18.7	39.5	5.9	34.4	98.7
ALT-n	8.4	0.7	31.2	5.1	16.3	6.1	20.2	13.9	17.8	38.5	6.1	35.7	100.6
ALT-o	7.8	0.6	32.2	5.2	16.0	6.0	19.7	13.0	18.2	38.8	6.0	36.4	101.9
ALT-p	7.0	0.6	31.7	8.2	16.4	6.9	21.3	16.4	16.8	34.8	6.7	33.1	88.7
ALT-q	7.6	0.6	31.7	7.3	16.7	7.1	18.9	13.2	19.0	39.1	6.1	32.5	93.2

The definitions that had the smallest Total Diff scores for gender and race were: ALT-c (VE+AR+MK), ALT-i (VE+AR), ALT-p (VE+MK), and ALT-q (VE+MK+GS). Subsequent analyses with Sample B will be limited to these four.

Sample B--FY84 Navy Recruit Population

ASVAB scores were collected for 78,193 FY84 Navy recruits.

Procedure

Percentage distributions of males and females across mental levels for the AFQT and the four alternatives were computed and listed in Table 8. Percentage distributions for Whites and Blacks are listed in Table 9.

Table 8

Percentage Distributions by Gender Across Mental Level for Navy Recruit Population FY84

	Mental Levels						
	I		II		IIIA		
Definitions	M	F	M	F	M	F	
AFQT	5.1	4.6	31.7	30.9	20.3	22.7	
ALT-c (VE+AR+MK)	5.3	3.2	34.4	29.3	19.6	21.5	
ALT-i (VE+AR)	6.1	4.4	33.8	30.0	18.1	19.1	
ALT-p (VE+MK)	4.6	2.9	30.4	28.7	18.9	21.6	
ALT-q (VE+MK+GS)	5.6	2.7	32.0	28.1	20.0	21.5	
	IIIB		IV		V		
Definitions	M	F	M	F	M	F	Total Diff
AFQT	30.4	36.4	12.5	5.4	0.0	0.0	16.8
ALT-c (VE+AR+MK)	28.3	26.8	12.4	9.2	0.1	0.0	13.9
ALT-i (VE+AR)	28.4	36.4	13.5	10.1	0.1	0.0	18.0
ALT-p (VE+MK)	29.8	35.1	16.2	11.6	0.2	0.1	16.1
ALT-q (VE+MK+GS)	24.4	30.5	17.8	17.1	0.2	0.1	15.2

Table 9

Percentage Distributions by Race Across Mental Level
for Navy Recruit Population FY84

	Mental Levels						
	I		II		IIIA		
Definitions	W	B	W	B	W	B	
AFQT	5.8	0.8	35.3	11.5	21.7	14.4	
ALT-c (VE+AR+MK)	5.8	0.8	37.8	12.4	20.8	14.2	
ALT-i (VE+AR)	6.9	0.9	37.3	12.1	19.2	12.8	
ALT-p (VE+MK)	5.0	0.8	33.6	11.7	20.2	13.7	
ALT-q (VE+MK+GS)	6.1	0.7	35.2	11.6	21.3	14.0	
	IIIB		IV		V		
Definitions	W	B	W	B	W	B	Total Diff
AFQT	28.6	44.4	8.6	28.9	0.0	0.0	72.2
ALT-c (VE+AR+MK)	26.9	41.6	8.6	30.9	0.1	0.1	74.0
ALT-i (VE+AR)	27.2	40.0	9.3	34.0	0.1	0.1	75.1
ALT-p (VE+MK)	28.9	38.5	12.2	35.1	0.2	0.2	65.1
ALT-q (VE+MK+GS)	23.9	31.1	13.3	42.1	0.2	0.4	72.5

Results and Discussion

As indicated by Table 8, the percent of males is greater than the percent of females in categories I, II, and IV for the AFQT and all four alternatives. These results are consistent with the results from Sample A. The Total Diff score for the AFQT is 16.8 percent while the scores for the four alternatives are relatively close. These scores ranged from 13.9 to 18.0 percent. Only ALT-c with a Diff score of 13.9 percent would result in a lower percentage difference. Although the other alternatives have larger percentage differences, the amount is minimal.

As indicated by Table 9, the percent of Whites is larger than the percent of Blacks in mental categories I, II, and IIIA, while the Blacks have larger percentages in categories IIIB and IV. This trend is similar to the AYP with the exception of category IIIB, where the Whites have a larger percentage. The Total Diff scores for the four alternatives range from 65.1 to 75.1 percent, with a score of 72.2 percent for the current AFQT. Adoption of any of these four alternatives would mean relatively little change from the current AFQT.

DISCUSSION

The alternatives found to have the least amount of percent change between males and females in Sample A were VE+AR+MK, VE+AR, VE+MK, and VE+MK+GS. These four alternative definitions also had "Total Diff" scores comparable to the current AFQT with regard to the percent of Whites and Blacks in Sample A. In terms of Sample B, the same four alternatives continued to show a minimum negative impact on females and Blacks.

As discussed earlier, Horne (1986) suggested that the length and composition of the alternatives be considered in evaluating for comparability to the AFQT. Therefore, in order to determine which one of the four alternatives should be recommended, length and composition were utilized as additional criteria.

In terms of length, Horne suggested that the length of the new AFQT definition be long rather than short. The current AFQT has a total of 130 scored items. However, the NO portion that has 50 items is halved, thus the scores range from 0 to 105 at half point increments. The number of items scored for the four alternative definitions are: (1) VE+MK (75 items), (2) VE+AR+MK (100 items), (3) VE+AR (80 items), and (4) VE+MK+GS (105 items). Based on length advantages of longer tests, the alternatives VE+AR and VE+MK would be too short because they both have significantly fewer items than the current AFQT.

With regard to composition, Horne (1986) proposed that the new AFQT definition be balanced in terms of verbal and math constructs. In addition, the new definition should have less weight on the technical construct. Of the four alternative AFQT definitions, only VE+AR+MK has a balance of verbal and math constructs. The VE component is composed of WK and PC thereby, giving the alternative definition two verbal constructs (WK and PC) and two math constructs (AR and MK). In terms of the technical construct, all four alternative AFQT definitions would qualify as none of them contain a technical subtest.

Finally, the Horne study found the following alternative AFQT definitions to be the most promising: (1) VE+AR+MK, (2) VE+AR+GS+MK, and (3) 2AR+2WK+2PC+MK. In another similar study by the Air force Human Resources Laboratory (AFHRL), Wegner and Ree (1986) recommended the following alternative AFQT definitions: (1) VE+AR+MK, (2) VE+AR, and (3) 2AR+2WK+2PC+MK. Of the four AFQT alternative definitions found suitable for replacement of the current AFQT, only VE+AR+MK was recommended by both Horne (1986) and Wegner and Ree (1986).

RECOMMENDATIONS

The alternative AFQT definition VE+AR+MK is the most promising replacement for the current AFQT and is recommended as the composite definition to replace the current AFQT. In comparison to the current AFQT, it had a comparable predictive validity and showed a minimum negative impact on females and Blacks in both Samples A and B. In addition, VE+AR+MK is: (1) adequately long enough to minimize test compromise, (2) composed of a balanced verbal and math construct, and (3) is recommended by two parallel studies. Therefore, the alternative VE+AR+MK is recommended.

REFERENCES

Earles, J. A., Giuliano, T., Ree, M. J., & Valentine, L. D. (1983). The 1980 youth population: an investigation of speeded subtests (Unpublished manuscript). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Horne, G. E. (1986). Evaluation of alternative compositions of the Armed Forces Qualification Test (AFQT) (Memorandum 86-111). Alexandria, VA: Center for Naval Analyses.

Maier, M. H., & Sims, W. H. (1982). Construction of an ASVAB score scale in the 1980 reference population (Memorandum 82-3288). Alexandria, VA: Center for Naval Analyses.

Maier, M. H., & Sims, W. H. (July 1986). The ASVAB score scales: 1980 and the World War II (CNR 116). Alexandria, VA: Center for Naval Analyses.

Sims, W. H., & Maier, M. H. (1983). The appropriateness for military applications of the ASVAB subtests and score scale in the new 1980 reference population (Memorandum 83-3102). Alexandria, VA: Center for Naval Analyses.

Wegner, T. G., & Ree, M. J. (1985). Armed Services Vocational Aptitude Battery: Correcting the speeded subtests for the 1980 youth population (AFHRL-TR-85-14). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory. (AD-A158 823)

Wegner, T. G., & Ree, M. J. (1986). Alternative Armed Forces Qualification Test composites (AFHRL-TP-86-27). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

APPENDIX A

VALIDITY OF ALTERNATIVE AFQT DEFINITIONS

Table A-1

Validity of Alternative AFQT Definitions for Schools
Using the General Technical Composite (VE+AR)

Course Number[a]	6001	6005	6125	6167	6301	6302	6477	
N of Cases	482	380	1,587	374	140	260	599	Avg
VE+AR	.33	.32	.46	.23	.50	.60	.11	.36
AFQT%	.33	.38	.46	.23	.51	.56	.11	.37
Raw AFQT	.32	.38	.45	.23	.51	.56	.11	.36
AFQT Alternatives[b]								
ALT-a	.33	.30	.46	.19	.46	.57	.13	.36
ALT-b	.31	.30	.44	.23	.47	.55	.11	.35
ALT-c	.34	.34	.48	.24	.53	.61	.13	.38
ALT-d	.29	.29	.47	.18	.49	.58	.13	.36
ALT-e	.33	.33	.48	.23	.53	.61	.14	.38
ALT-f	.32	.32	.48	.18	.49	.61	.11	.37
ALT-g	.33	.29	.44	.23	.48	.55	.15	.35
ALT-h	.30	.27	.45	.16	.36	.53	.11	.33
ALT-i	.31	.30	.46	.18	.51	.59	.13	.36
ALT-j	.33	.30	.48	.21	.43	.56	.16	.37
ALT-k	.32	.33	.49	.23	.51	.61	.14	.38
ALT-l	.34	.32	.49	.23	.48	.59	.13	.38
ALT-m	.33	.32	.49	.23	.47	.58	.15	.38
ALT-n	.31	.30	.48	.22	.47	.58	.16	.37
ALT-o	.32	.31	.48	.22	.45	.58	.16	.37
ALT-p	.30	.30	.42	.24	.53	.56	.16	.35
ALT-q	.28	.27	.44	.22	.50	.55	.15	.35

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

Table A-2

Validity of Alternative AFQT Definitions for Schools
Using the Mechanical Composite (VE+MC+AS)

Course Number[a]	6025	6119	6120	6286	6292	
N of Cases	103	391	300	203	181	Avg
VE+MC+AS	.48	.34	.29	.43	.22	.34
AFQT%	.45	.32	.34	.32	.22	.32
Raw AFQT	.46	.32	.34	.32	.22	.32
AFQT Alternatives[b]						
ALT-a	.50	.35	.35	.42	.27	.36
ALT-b	.46	.39	.28	.28	.28	.33
ALT-c	.46	.37	.35	.38	.30	.36
ALT-d	.50	.38	.30	.31	.23	.34
ALT-e	.47	.36	.35	.37	.28	.36
ALT-f	.49	.37	.33	.36	.27	.35
ALT-g	.48	.38	.34	.39	.27	.36
ALT-h	.49	.37	.31	.38	.28	.35
ALT-i	.47	.34	.33	.34	.24	.33
ALT-j	.53	.43	.36	.43	.30	.40
ALT-k	.49	.41	.33	.35	.29	.37
ALT-l	.51	.41	.35	.41	.30	.39
ALT-m	.52	.42	.35	.41	.29	.39
ALT-n	.52	.43	.36	.39	.28	.39
ALT-o	.52	.43	.36	.38	.29	.39
ALT-p	.44	.37	.32	.29	.24	.33
ALT-q	.48	.40	.29	.27	.24	.33

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

Table A-3

Validity of Alternative AFQT Definitions for Schools
Using the Electronics Composite (AR+MK+EI+GS)

Course Number[a]	6027	6108	6131	6146	6376	6377	
N of Cases	172	135	118	189	117	245	Avg
AR+MK+EI+GS	.49	.39	.21	.26	.45	.29	.34
AFQT%	.40	.32	.26	.20	.27	.27	.29
RAW AFQT	.40	.32	.27	.21	.28	.27	.29
AFQT Alternatives[b]							
ALT-a	.44	.39	.26	.29	.36	.33	.35
ALT-b	.43	.41	.28	.31	.46	.32	.36
ALT-c	.54	.46	.38	.29	.40	.31	.39
ALT-d	.41	.36	.16	.24	.27	.29	.29
ALT-e	.50	.43	.33	.27	.35	.29	.35
ALT-f	.41	.36	.24	.26	.29	.29	.31
ALT-g	.52	.45	.33	.32	.47	.36	.40
ALT-h	.36	.35	.16	.29	.31	.33	.31
ALT-i	.43	.37	.25	.22	.25	.26	.30
ALT-j	.50	.44	.26	.35	.44	.36	.39
ALT-k	.50	.44	.30	.30	.40	.32	.37
ALT-l	.50	.44	.30	.34	.45	.36	.40
ALT-m	.52	.43	.29	.35	.44	.35	.39
ALT-n	.51	.44	.26	.33	.39	.34	.38
ALT-o	.51	.45	.26	.34	.40	.34	.38
ALT-p	.52	.46	.32	.26	.39	.29	.36
ALT-q	.47	.42	.22	.27	.37	.31	.34

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

Table A-4

Validity of Alternative AFQT Definitions for Schools Using the Clerical Composite (VE+NO+CS)

Course Number[a]	6020	
N of Cases	107	Avg
VE+NO+CS	.23	.23
AFQT%	.25	.25
Raw AFQT	.25	.25
AFQT Alternatives[b]		
ALT-a	.28	.28
ALT-b	.17	.17
ALT-c	.24	.24
ALT-d	.25	.25
ALT-e	.25	.25
ALT-f	.25	.25
ALT-g	.25	.25
ALT-h	.20	.20
ALT-i	.26	.26
ALT-j	.22	.22
ALT-k	.24	.24
ALT-l	.25	.25
ALT-m	.24	.24
ALT-n	.21	.21
ALT-o	.21	.21
ALT-p	.22	.22
ALT-q	.21	.21

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

Table A-5

Validity of Alternative AFQT Definitions for Schools
Using the BE/E Composite (AR+2MK+GS)

Course Number[a]	6036	6070	6093	6289	
N of Cases	219	370	204	126	Avg
AR+2MK_GS	.35	.32	.23	.26	.30
AFQT%	.40	.31	.26	.31	.32
Raw AFQT	.39	.31	.27	.31	.32
AFQT Alternatives[b]					
ALT-a	.44	.31	.30	.34	.34
ALT-b	.31	.33	.28	.21	.30
ALT-c	.38	.31	.27	.29	.32
ALT-d	.38	.29	.32	.27	.32
ALT-e	.39	.30	.27	.30	.31
ALT-f	.41	.32	.29	.28	.33
ALT-g	.39	.29	.31	.32	.32
ALT-h	.42	.33	.34	.29	.35
ALT-i	.39	.27	.27	.29	.30
ALT-j	.43	.34	.35	.32	.36
ALT-k	.39	.32	.31	.28	.33
ALT-l	.42	.33	.33	.32	.35
ALT-m	.42	.35	.33	.31	.36
ALT-n	.40	.33	.35	.27	.34
ALT-o	.41	.34	.35	.29	.35
ALT-p	.32	.25	.27	.26	.27
ALT-q	.33	.27	.32	.24	.29

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

Table A-6

Validity of Alternative AFQT Definitions for Schools Using the Repairman Composite (AR+MC+AS)

Course Number[a]	6068	
N of Cases	202	Avg
VE+NO+CS	.40	.40
AFQT%	.13	.13
Raw AFQT	.14	.14
AFQT Alternatives[b]		
ALT-a	.24	.24
ALT-b	.25	.25
ALT-d	.11	.11
ALT-e	.18	.18
ALT-f	.15	.15
ALT-g	.28	.28
ALT-h	.25	.25
ALT-i	.12	.12
ALT-j	.31	.31
ALT-k	.19	.19
ALT-l	.26	.26
ALT-m	.28	.28
ALT-n	.24	.24
ALT-o	.26	.26
ALT-p	.16	.16
ALT-q	.15	.15

[a]See Table 2 for definitions.

[b]See Table 1 for definitions.

DISTRIBUTION LIST

Naval Military Personnel Command (NMPC-481)
Chief of Naval Operations (OP-135L)
Office of Assistant Secretary of Defense (FM&P)
Defense Technical Information Center (DTIC) (2)